

Distribution of Centromere-Like *parS* Sites in Bacteria: Insights from Comparative Genomics^{∇†}

Jonathan Livny, Yoshiharu Yamaichi, and Matthew K. Waldor*

Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, and Howard Hughes Medical Institute, 181 Longwood Avenue, Boston Massachusetts 02115

Received 31 July 2007/Accepted 18 September 2007

Partitioning of low-copy-number plasmids to daughter cells often depends on ParA and ParB proteins acting on centromere-like *parS* sites. Similar chromosome-encoded *par* loci likely also contribute to chromosome segregation. Here, we used bioinformatic approaches to search for chromosomal *parS* sites in 400 prokaryotic genomes. Although the consensus sequence matrix used to search for *parS* sites was derived from two gram-positive species, putative *parS* sites were identified on the chromosomes of 69% of strains from all branches of bacteria. Strains that were not found to contain *parS* sites clustered among relatively few branches of the prokaryotic evolutionary tree. In the vast majority of cases, *parS* sites were identified in origin-proximal regions of chromosomes. The widespread conservation of *parS* sites across diverse bacteria suggests that *par* loci evolved very early in the evolution of bacterial chromosomes and that the absence of *parS*, *parA*, and/or *parB* in certain strains likely reflects the loss of one or more of these loci much later in evolution. Moreover, the highly conserved origin-proximal position of *parS* suggests *par* loci are primarily devoted to regulating processes that involve the origin region of bacterial chromosomes. In species containing multiple chromosomes, the *parS* sites found on secondary chromosomes diverge significantly from those found on their primary chromosomes, suggesting that chromosome segregation of multipartite genomes requires distinct replicon-specific *par* loci. Furthermore, *parS* sites on secondary chromosomes are not well conserved among different species, suggesting that the evolutionary histories of secondary chromosomes are more diverse than those of primary chromosomes.

Dividing cells have mechanisms to ensure that their genetic material is faithfully segregated to daughter cells. Eukaryotes utilize a conserved mitotic apparatus in which a variety of proteins act at particular DNA sites known as centromeres to direct chromosome segregation. The mechanisms that account for chromosome and plasmid segregation in prokaryotes are less understood. Partitioning (*par*) genes are known to be critical for the stable inheritance of several low-copy-number plasmids (14), and in some cases it is now clear that Par proteins mediate the active partitioning of duplicated plasmids to daughter cells (14, 26). Many bacterial chromosomes encode orthologues of plasmid Par proteins (21), but with few exceptions, the role of these proteins in the segregation of duplicated chromosomes to daughter cells is not known.

Plasmid-encoded *par* loci consist of two genes, often called *parA* and *parB*, and a *cis*-acting centromere-like site, often referred to as *parS*. ParB proteins bind to cognate *parS* sites, forming a nucleoprotein complex. ParA proteins are ATPases that, in a few cases, have been shown to form dynamic filaments (3, 14, 19, 24, 36, 37, 44). ParA proteins interact with ParB/*parS* complexes and are, like *parB* and *parS*, essential for plasmid partitioning. Recent elegant *in vitro* reconstitution

studies strongly suggest that ParA, ParB, and *parS* are the key components of plasmid partitioning systems (20).

To date the function of chromosomal *par* genes is not as well defined. While chromosomal *par* loci appear to contribute to chromosome localization and segregation (16, 22, 28, 30, 32, 34, 52), there is increasing evidence that they are not essential for accurately partitioning chromosomes to daughter cells, perhaps due to redundancy in the mechanisms that account for chromosome partitioning. Chromosomal *parAB* loci are usually found in the origin-proximal regions of chromosomes. In *Bacillus subtilis* and *Vibrio cholerae*, *par* loci have been shown to contribute to origin localization (16, 34, 35, 51). In *B. subtilis*, ParB (Spo0J) is implicated in the control of initiation of chromosome replication as well (34, 35, 46, 57). *par* loci also have specialized roles in certain bacteria. For example, in *B. subtilis*, Par homologues regulate entry into sporulation (11, 28, 49), and in *Caulobacter crescentus*, the ParB/*parS* complex influences cell division (42, 53).

Phylogenetic analyses have revealed that chromosome-encoded ParA and ParB proteins cluster into a subgroup that is distinct from plasmid-encoded Par proteins (13, 21, 26, 62). The chromosomal subgroup of Par proteins includes proteins from both gram-positive and gram-negative bacteria. Despite the conservation of chromosome-encoded ParA and ParB proteins from diverse bacteria, not all bacterial species contain Par homologues. For example, several well-studied *Gammaproteobacteria*, including *Escherichia coli*, *Salmonella* sp., *Haemophilus* sp., and *Yersinia* sp., lack chromosomal *par* genes. Interestingly, in bacteria that have complex genomes consisting of more than one chromosome, the Par proteins encoded on the smaller chromosome(s) tend to cluster in phylogenetic

* Corresponding author. Mailing address: Channing Laboratory, Brigham and Women's Hospital, 181 Longwood Avenue, Boston MA 02115. Phone: (617) 525-4646. Fax: (617) 525-4660. E-mail: mwaldor@rics.bwh.harvard.edu.

† Supplemental material for this article may be found at <http://jb.asm.org/>.

[∇] Published ahead of print on 28 September 2007.

trees with plasmid-encoded Par proteins (13, 21, 62), which are more diverse than chromosome-encoded proteins.

The *cis*-acting *parS* sites in plasmid *par* loci are located close to the *parAB* genes. The sequences and structures of plasmidic *parS* sequences are highly variable and often complex. For example, in the F plasmid, *parS* (*sopC*) consists of 12 tandem repeats of a 43-bp sequence (6). In P1, as in F, *parS* is found downstream of *parB*, but this site consists of a single ~80-bp sequence that includes two ParB binding sites flanking a binding site for integration host factor (17, 25). The binding of ParB and other partition factors to *parS* sites likely induces functionally significant topological changes in these DNA sequences (6, 25, 26, 58).

Chromosomal *parS* sites were first described in *Bacillus subtilis* by Lin and Grossman (38). They identified eight *B. subtilis* *parS* sites bound by Spo0J *in vivo* with a chromatin immunoprecipitation assay. All of these sites were located in the origin-proximal 20% of the *B. subtilis* chromosome and consisted of a similar 16-bp sequence that included an imperfect 8-bp inverted repeat. Using a consensus Spo0J binding sequence of 5'-TGTTNCACGTGAAACA-3', Lin and Grossman also identified potential *parS* sites in 10 genomes in the relatively small genome database that was available at that time. Since that time, chromosomal *parS* sites have been experimentally identified in seven other bacterial species (4, 13, 22, 30, 33, 42, 43, 60). In nearly all cases, these chromosomal *parS* sites are very similar to the *B. subtilis* consensus sequence in structure, length, and sequence.

Although most prokaryotic genomes are composed of a single chromosome, it is now clear that the genomes in several different families of prokaryotes contain multiple chromosomes (15, 31). In bacteria with complex genomes comprised of more than one chromosome, the largest (primary) chromosome usually contains the majority of essential genes, and the smaller (secondary) chromosome(s) contains relatively few essential genes (15, 31). There is relatively little knowledge of *par* loci in bacteria with complex genomes. *parS* sites have been experimentally identified in *Vibrio cholerae* and *Burkholderia cenocepacia*, bacterial species whose genomes are comprised of two and three chromosomes, respectively. In both organisms, the *parS* sequences on the large chromosome are nearly identical to the *B. subtilis* site, whereas the *parS* sites on the secondary chromosomes differ significantly from the *B. subtilis* consensus sequence (13, 60).

Here we used bioinformatic approaches to search for putative *parS* sites in all the sequenced replicons, including all chromosomes and extrachromosomal elements, available in the NCBI database. We found that 69% of strains contain putative chromosomal *parS* sites and that species bearing putative *parS* sites are found in all branches of prokaryotes. In the vast majority of cases, *parS* sites were identified in the origin-proximal region of the chromosome relatively close to the *parAB* loci. Remarkably, no *parS* sites characteristic of primary chromosomes were identified on secondary chromosomes. However, we identified distinct family-specific sets of *parS* sites on the second and third chromosomes (referred to as *parS2* and *parS3* sites, respectively) of most bacterial species with complex genomes. The *parS2* and *parS3* sites were also found in the origin-proximal region of the chromosome. With one exception, when *parS2* and *parS3* sites were identified on sec-

ondary chromosomes, they were not found on primary chromosomes of species with multipartite genomes. Overall, our observations suggest that *par* loci are primarily devoted to regulating processes that involve the origin region of bacterial chromosomes. Furthermore, bacteria harboring multiple replicons appear to require distinct replicon-specific *par* loci.

MATERIALS AND METHODS

Programs and search parameters used. Blast comparisons were conducted using BLASTN and BLASTP 2.0MP-WashU (1). Search parameters B (the maximum number of database sequences for which any alignments will be reported) and V (the maximum number of one-line descriptions of significant database sequences reported) were set to 10,000. Unless otherwise noted, all other search parameters were set to default values. Motif searches were conducted using Patser v3e.1 (54) and RNAMotif v3.0.4 (41). For Patser searches, the a priori nucleotide probabilities used to convert the alignment matrix to a weight matrix were set to 1 for all four nucleotides. Scrambled matrices were created by shuffling the columns in each half-site of the *parS* consensus matrix symmetrically so that the resulting matrices maintained the same palindromic structure, length, and overall base frequencies as the *parS* matrix but corresponded to different primary consensus sequences.

Sequence databases. Genome sequence files (.fna extensions), protein sequence files (.faa extensions), and open reading frame (ORF) annotation files (.ptt extensions) were obtained from the NCBI ftp database (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/).

***oriC* predictions.** In circular bacterial chromosomes, the leading strand often contains more G than C nucleotides, and the origin of replication can be identified using this G/C skew (40). In many but not all published genomes, the annotation ± 1 has been assigned based on the location of the GC skew minimum (and thus likely corresponds to *oriC*). Another feature common to the *oriC* regions of diverse species is their proximity to genes encoding homologues of the replication initiator protein DnaA and of the glucose-inhibited division protein GidA (9, 18, 23, 27, 45, 48, 63). For our analyses we defined the putative *oriC*s as follows.

(i) **Circular primary chromosomes.** If the annotated +1 was within 2% of the genome size from both the putative DnaA- and GidA-encoding genes (identified by homology to the DnaA and GidA of *V. cholerae*, *Escherichia coli*, and *B. subtilis*), the putative *oriC* was assigned at the putative +1. Alternatively, if the annotated +1 was farther than 2% of the genome size from either of the putative DnaA- and GidA-encoding genes, the GC minimum was determined using GenSkew (<http://mips.gsf.de/services/analysis/gen skew>) and the putative *oriC* was assigned based on this minimum.

(ii) **Linear primary chromosomes.** The *oriC* was assigned at a position directly upstream of the *dnaA* gene.

(iii) **Secondary chromosomes.** No DnaA or GidA homologues were identified on secondary chromosomes. Thus, the *oriC*s of secondary chromosomes were assigned based on the GC skew minima identified by GenSkew.

Construction of *parS2/IR* consensus matrices. The *V. cholerae* chromosome II *parS* site consensus matrix was created using the nine *V. cholerae* *parS2* sites identified by Yamaichi et al. (60). The *B. cenocepacia* chromosome II and chromosome III *parS* site consensus matrices were created using the six and the four *B. cenocepacia* *parS2* and *parS3* sites, respectively, that were identified by Dubarry et al. (13). To construct the *Ralstonia parS2* consensus matrix, RNAMotif was used to search the sequence of chromosome II of *Ralstonia eutropha* H16 for perfect palindromes corresponding to the motif 5'-TTN(4)CGN(4)AA-3'. This motif was based on the putative *parS* site identified by Dubarry et al. on pGM1000MP of *R. solanacearum*. The six *R. eutropha* sites identified in this search were incorporated into the *R. eutropha parS2* consensus matrix. To construct the consensus matrix for the secondary chromosome IR (*parS*) sites of *Brucella suis*, RNAMotif was used to search for any 7-bp inverted repeat flanking two central bases in the 6-kb region of the *Brucella suis* chromosome II that is centered at the *repABC* operon. This search led to the identification of one site that, similar to other IR sites, contained a central GC motif and was located between the *repA* and *repB* genes. Patser was then used to search for other sites similar to this sequence in the entire chromosome, leading to the identification of another palindromic sequence with a central GC directly upstream of the *repA* gene. These two sites were incorporated into the consensus matrix.

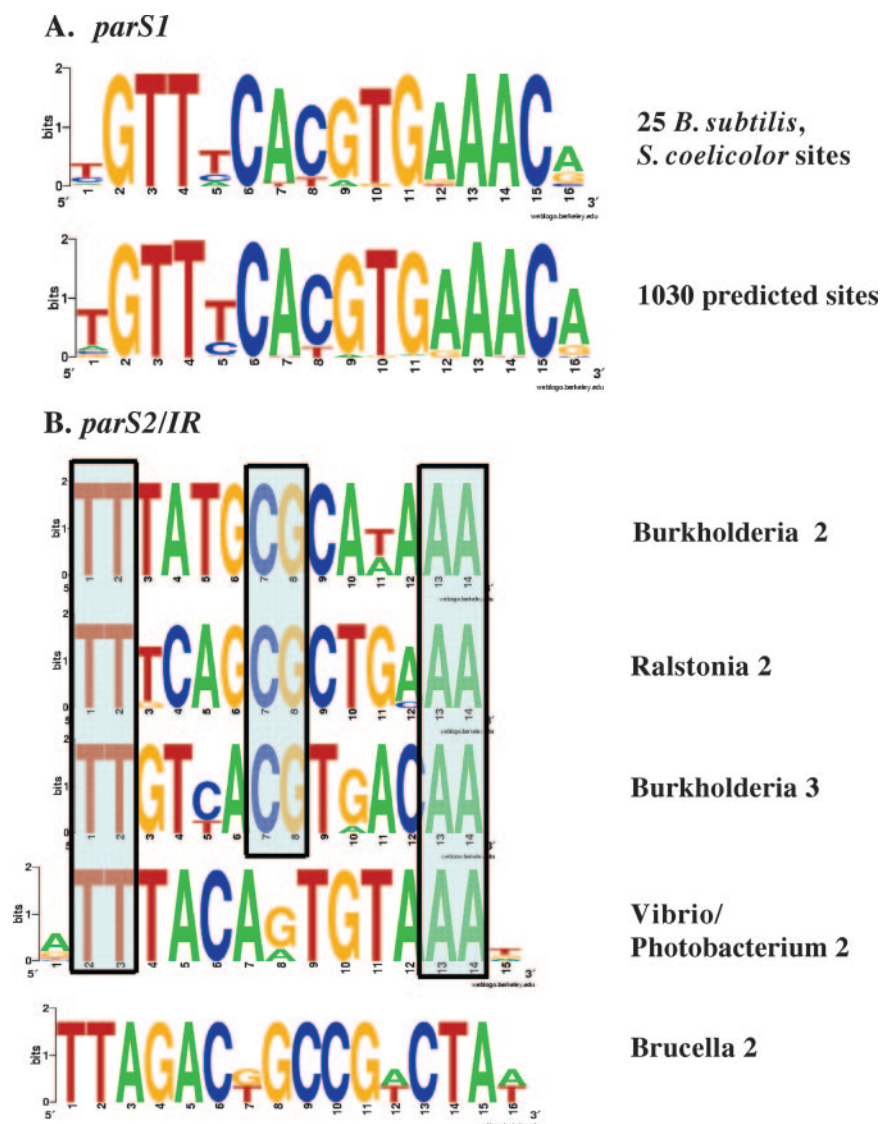


FIG. 1. WebLogos of putative *parS* site consensus sequences. A. Logos representing the 25 *parS1* sites used to construct the *S. coelicolor* and *B. subtilis* consensus matrix and the 1,030 putative *parS* sites identified using this matrix. B. Logos representing the putative *parS* sites identified in secondary chromosomes using the indicated consensus matrices. The shaded boxes highlight conserved motifs.

RESULTS AND DISCUSSION

Informatic approach used to search for putative *parS* sites in prokaryotic replicons. To search for putative *parS* sites in all 706 sequenced prokaryotic chromosomes, plasmids, and extra-chromosomal elements in the NCBI database, we generated a *parS* consensus matrix using previously characterized chromosomal *parS* sequences (Fig. 1A). Prior work had demonstrated that consensus matrices are a robust tool for identification of DNA sequence motifs in genomes (50). To help ensure the accuracy of our predictions, only *parS* sites that have been shown to bind ParB in vivo were used to create the consensus matrix. These included 15 sites from *Streptomyces coelicolor* (30) and 10 sites from *Bacillus subtilis* (8, 38). The program Patser (54) was then used to search for sequences corresponding to this matrix in all 706 prokaryotic replicons in the NCBI database (see Table S1 in the supplemental material).

Patser searches were repeated with different minimum score thresholds, and the optimal minimum score was empirically found to be 15, based on the concordance of the predicted *parS* sites with the results of previous studies. With the minimum score set to 15, no *parS* sites were identified in *Haemophilus influenzae* and *E. coli*, species that are known not to encode Par proteins. In addition, using this threshold we identified all previously characterized *parS* sites in *Pseudomonas putida* and *Caulobacter crescentus* and on the primary chromosomes of *Burkholderia cenocepacia* and *Vibrio cholerae* (13, 22, 42, 43, 60), even though the *parS* sequences from these organisms were not used in the construction of the *parS* consensus matrix. Finally, no additional *parS* sites were identified in these replicons, including several putative *parS* sites in *B. subtilis* and *V. cholerae* that were identified by bioinformatics but shown to not be functional (38, 60).

Although our search parameters appear to have a high degree of specificity, the sensitivity of our search was not perfect, as several previously identified *parS* sites were not detected. For example, 2 of the 10 verified *B. subtilis* *parS* sites were missed in this search. Even though both of these recently described sites (8) were included in the consensus matrix, their sequences diverge significantly from the sequences of the other 23 sites used to generate the matrix, and their identification in a Patser search required a threshold score below 15. These findings suggest that some bona fide *parS* sites in other replicons whose sequences diverge from those of most other sites have been missed in our predictions.

In *Pseudomonas aeruginosa*, our search detected 4 of the 10 putative *parS* sites annotated by Bartosik et al. (4). The four sites we identified diverge least from the *B. subtilis* *parS* consensus sequence, are closest to the *P. aeruginosa* origin of replication, and are the sites that Bartosik et al. suggested were most likely to be functional (4). It is not known if the six putative *P. aeruginosa* *parS* sites missed in our search represent functional *parS* sites. Thus, some of the previously identified *parS* sites that were missed in our predictions may not correspond to functional *parS* sites.

The two putative *parS* sites we identified in *Helicobacter pylori* were not the same two sites previously reported by Lee et al. (33). The sites we found in our search are the same length as and diverge by only a single base from the *B. subtilis* consensus sequence; in contrast, the sites identified by Lee et al. are either 1 base longer or shorter than the *B. subtilis* consensus sequence and diverge from this sequence by >3 bases. Also, the putative *parS* sites we identified are much closer to the *H. pylori* origin of replication than the sites found by Lee et al. (33). As discussed below, the proximity of the two *parS* sites we identified to the putative *H. pylori* origin of replication supports their validity.

***parS* sites are found in diverse prokaryotes and only on primary chromosomes.** Our search resulted in the identification of 1,030 putative *parS* sites in 276 (69%) of the 400 sequenced strains in the NCBI database (see Table S2 in the supplemental material). Remarkably, although the *parS* consensus matrix was derived from two gram-positive species, putative *parS* sites were found in all branches of the prokaryotic evolutionary tree, including in four strains of archaea (Fig. 2), and their sequences were found to be very well conserved (Fig. 1A). The widespread conservation of this site across diverse species suggests that *parS* and probably *par*-based segregation systems arose very early in the evolution of prokaryotes. For the most part, with the exception of two branches of *Gamma-proteobacteria*, organisms lacking *parS* sites were scattered throughout the prokaryotic evolutionary tree (Fig. 2). Surprisingly, of the 1,030 *parS* sites identified, only 1 was identified on a plasmid and none were identified on secondary chromosomes, despite the fact that plasmids and secondary chromosomes comprise 7.2% of all sequences in the NCBI database. Thus, the *parS* sites identified in this search appear to be distinctive features of primary chromosomes.

Variability in number but not location of *parS* sites in different species. As shown in Fig. 3A, there is considerable variability in the number of predicted *parS* sites in different strains. While the majority of the *parS*⁺ strains contain one to four putative *parS* sites, 23 strains belonging to 11 genera are

predicted to encode eight or more *parS* sites. In some cases, all species in the genera (e.g., *Streptomyces* spp.) have a large number of predicted *parS* sites; in other cases, there was a large range in the number of *parS* sites in different species of the same genera (e.g., *Lactobacillus* spp.) (Fig. 2). The number of putative *parS* sites predicted in each chromosome did not correlate with chromosome size. For example, the 3-Mbp chromosome of *Listeria innocua* contains 20 putative *parS* sites, while the 7-Mbp chromosome of *Pseudomonas fluorescens* Pf-5 contains only 2. Indeed, among the 400 strains analyzed, the number of *parS* sites predicted per Mbp of chromosome ranges nearly 42-fold, from 0.2 in *Acidobacteria bacterium* Ellin345 to 7.5 in *Listeria welshimeri*.

As shown in Fig. 4A, the locations of the 1,029 predicted chromosomal *parS* sites are not distributed randomly throughout the respective genomes. The vast majority of the putative sites were identified within origin-proximal regions of their respective chromosomes. More than 92% of the sites were located in chromosomal regions corresponding to 15% of the respective replicon's size centered at its *oriC* (referred to below as the "15% *oriC* region") (Fig. 4A). The percentage of functional *parS* sites within the 15% *oriC* region is likely to be even higher, since many of the 82 sites predicted outside this region are found in species that lack *parAB* genes. The average distance of the predicted *parS* sites from the respective *oriC* was 2.6% of the replicon size. Even in replicons encoding 10 or more *parS* sites, all sites are clustered within this 15% *oriC* region.

To determine if this positional bias was specific to the *parS* consensus sequence rather than to 16 bp palindromes in general, we repeated our search using two different "scrambled" matrices representing palindromic motifs of the same length as the *parS* consensus but corresponding to different primary sequences (see Materials and Methods). Searches using these palindromes yielded only 21 and 91 sites, respectively, suggesting that the *parS* consensus sequence is greatly overrepresented in prokaryotic chromosomes compared to similar palindromic motifs. Moreover, only 18% of the sites corresponding to the scrambled matrices were found within the 15% *oriC* regions, suggesting that the proximity of putative chromosomal *parS* sites to the respective *oriCs* is specific to the *parS* consensus sequence. The remarkable positional conservation of *parS* sites suggests that the function of chromosomal *par* loci is highly dependent on the proximity of *parS* to *oriC*.

Most primary chromosomes that contain *parS* sites also encode ParA and ParB homologues. To identify ParA and ParB homologues on primary chromosomes, a database of the 1,151,128 annotated proteins in the NCBI database was compared by BLAST to the ParA and ParB proteins encoded by *B. subtilis* and by the primary chromosome of *V. cholerae*. The minimum BLAST score was set to 350 to eliminate false positives in species known not to encode true ParA and ParB homologues, such as *E. coli* and *H. influenza*. In total, 255 and 282 ParA and ParB homologues, respectively, were identified. All of the putative ParA and ParB homologues identified were encoded on primary chromosomes, an observation consistent with prior studies that suggested chromosomal Par proteins are phylogenetically distinct from Par proteins encoded by secondary chromosomes and plasmids (13, 21). Two hundred forty-

strains (5%) encode ParA and ParB homologues but no predicted *parS* sites (ParAB strains).

Proximity of *parS* sites to *parAB* loci on primary chromosomes. In all previously characterized chromosome-borne *par* loci, the *parA* and *parB* genes are located relatively close to one or more *parS* sites; furthermore, with the exception of *H. pylori* (33), the reported *par* genes and *parS* sites are usually found in close proximity to *oriC*. In the majority of Par3 strains we identified, these positional relationships between *parA*, *parB*, *parS*, and *oriC* are conserved. First, in 197 (88%) of the 224 Par3 strains, the distance between *parA* or *parB* and the nearest *parS* site is less than 10% of the chromosome size (denoted as ParABS strains). Second, in 172 (87%) of the 197 ParABS strains, the putative *parA*, *parB*, and nearest *parS* sites are all located less than 10% of the chromosome size away from the putative *oriC*. In the other 25 Par3 strains, the *parAB* genes are located near each other but farther than 10% of the chromosome size away from the nearest *parS* site (ParAB-S strains) (Fig. 1). These include all five strains of *Helicobacter* sp. in the database, in which all predicted *parS* sites are found within their respective 3% *oriC* regions but more than 20% of the genome size away from the *parAB* loci. The conserved positional relationships of *parA*, *parB*, *parS*, and *oriC* in Par3 strains suggest that these genes and sites have coevolved and that there is selective pressure promoting their proximity.

Phylogenetic analysis of the ParABS system in primary chromosomes. We next examined whether the number and relative chromosomal locations of the different putative ParABS components encoded by the 400 primary chromosomes in the NCBI database correlate with the phylogenetic relationships of these strains (Fig. 2). In most cases, the numbers of *parS* sites found in different strains of the same species vary by no more than one or two sites (Fig. 2). However, in a few genera, such as *Lactobacillus*, the number of putative *parS* sites predicted per strain varied by more than 10.

Our analyses also revealed that different ParABS profiles tend to cluster among phylogenetically related species, genera, and, in many cases, classes and phyla (Fig. 2). Many of the Par0 strains are clustered in two branches of *Gammaproteobacteria* that include *Escherichia*, *Yersinia*, *Salmonella*, *Buchnera*, and *Haemophilus* genera and in the *Mesoplasma*/*Mycoplasma*/*Ureaplasma*/*Candidatus* branches of the *Firmicute* class. Moreover, all but 2 of the 37 ParBS strains are in the closely related *Staphylococcus*, *Streptococcus*, and *Lactococcus* genera. Finally, most of ParAB-S strains belong to several closely related species of *Bacteroides* and *Gammaproteobacteria*. There are some notable exceptions to the correlation between ParABS profiles and phylogeny (Fig. 2). For example, while *Anaplasma phagocytophilum* is a Par0 strain, *Anaplasma marginale* is a Par3 strain. Thus, in some species loss of the Par loci appears to have occurred relatively recently in their evolution.

Secondary chromosomes do not have the *parS* sequences found on primary chromosomes. ParABS systems encoded by the secondary chromosomes of *B. cenocepacia* and *V. cholerae* have been described (13, 60). However, the *parS* sites present on chromosomes II and III in *B. cenocepacia* and on *V. cholerae* chromosome II differ significantly in primary sequence and, in the case of *V. cholerae*, structure, from *parS* sites on primary chromosomes (13, 60). Thus, it was not surprising that these previously identified *parS2* and *parS3* sites were not iden-

tified in the search described above. However, it was surprising that of the over 1,000 putative *parS1* sites identified and discussed above, none were found on secondary chromosomes, despite the fact that these replicons comprise 4.9% of the total DNA sequence in the NCBI database. This underrepresentation of *parS1* sites on secondary chromosomes was not observed when the search was repeated using the scrambled *parS* matrices; in these searches, 8 of the 112 (7%) predicted sites were found on secondary chromosomes. The complete absence of *parS1* sites from secondary chromosomes suggests that there is selective pressure against the occurrence of such sequences on secondary chromosomes.

Identification of putative *parS* sites on secondary chromosomes. Thirty-seven replicons in the NCBI database are annotated as secondary chromosomes, including five replicons annotated as chromosome III and two annotated as linear chromosomes (Table 1). For our analyses, the 2.1-Mbp pGMI1000MP plasmid of *Ralstonia solanacearum* was added to this list, since all other *Ralstonia* spp. strains carry a second chromosome of approximately the same size as pGMI1000MP. Unlike *parS* sites on primary chromosomes, the characterized *parS* sites on the secondary chromosomes of *B. cenocepacia* and *V. cholerae* vary significantly both in primary sequence and in structure. The *V. cholerae parS2* sites are 15-bp sequences composed of 7-nucleotide inverted repeats separated by a central base (60), while the *parS* sites on *B. cenocepacia* chromosome II and chromosome III and the predicted *Ralstonia solanacearum parS2* site are composed of 14-bp palindromic sequences (13). Although the predicted *R. solanacearum parS2* and the *B. cenocepacia parS2* and *parS3* sequences share similar structures, their primary sequences diverge significantly (13). No putative *parS* sites on the secondary chromosomes of *Brucella* spp. have previously been described. However, these replicons are known to encode homologues of the RepABC proteins that mediate segregation of several alphaproteobacterial plasmids (5, 39) and of the secondary (linear) chromosomes of *Agrobacterium tumefaciens* (29). Like ParB, RepB binds palindromic sites, known as IR sites, that are located in very close proximity to *oriC* (5, 12, 39). To identify putative *parS* sites on secondary chromosomes, we constructed five separate matrices using the *parS*/IR sites from chromosome II of *V. cholerae*, *B. cenocepacia*, *R. eutropha*, and *Brucella suis* and from chromosome III of *B. cenocepacia* (see Materials and Methods). Patser was then used to search for sites corresponding to these consensus matrices in all 706 replicons in the NCBI database. The results of these searches are summarized in Table 1 and described below (see also Table S3 in the supplemental material).

Sequence homology, number, and distribution of *parS* sites on secondary chromosomes. Using the five consensus sequences described above, a total of 151 putative *parS* sites were identified on 28 of the 38 secondary chromosomes in the NCBI database (Table 1). In contrast with *parS1* sites, which are highly conserved among diverse species (Fig. 1A), the sequences of *parS*/IR sites found on secondary chromosomes were family specific; for example, the search using the *V. cholerae parS2* consensus matrix identified *parS2* sites on all the second chromosomes of the *Vibrionaceae*/*Photobacteriaceae* species, but not on the second chromosomes in species outside of this family. Overall, there is significant divergence of the

TABLE 1. *parS*/IR sites and ParAB/RepAB homologues predicted on secondary chromosomes

Strain	Chromosome	No. of <i>parS</i> /IR sites ^a	<i>par</i> site consensus ^b	Par/RepA score ^c	Par/RepB score ^c	Par/RepAB homologue ^b
<i>Brucella abortus</i> 9-941	II	2	Bs2	2,029	1,645	Bs2
<i>Brucella melitensis</i>	II	2	Bs2	2,029	1,497	Bs2
<i>Brucella melitensis</i> bv. abortus	II	2	Bs2	2,029	1,645	Bs2
<i>Brucella suis</i> 1330	II	2	Bs2	2,035	1,645	Bs2
<i>Burkholderia</i> 383	II	6	Bc2	1,092	1,753	Bc2
<i>Burkholderia</i> 383	III	6	Bc3	363	382	Bc2
<i>Burkholderia cenocepacia</i> AU 1054	II	6	Bc2	1,098	1,794	Bc2
<i>Burkholderia cenocepacia</i> AU 1054	III	6	Bc3	362	389	Bc2
<i>Burkholderia cenocepacia</i> HI2424	II	6	Bc2	1,098	1,794	Bc2
<i>Burkholderia cenocepacia</i> HI2424	III	6	Bc3	362	390	Bc2
<i>Burkholderia cepacia</i> AMMD	II	5	Bc2	1,069	1,731	Bc2
<i>Burkholderia cepacia</i> AMMD	III	7	Bc3	363	382	Bc2
<i>Burkholderia mallei</i> ATCC 23344	II	4	Bc2	1,047	1,495	Bc2
<i>Burkholderia pseudomallei</i> 1710b	II	4	Bc2	1,047	1,495	Bc2
<i>Burkholderia pseudomallei</i> K96243	II	4	Bc2	1,047	1,495	Bc2
<i>Burkholderia thailandensis</i> E264	II	4	Bc2	1,052	1,511	Bc2
<i>Burkholderia xenovorans</i> LB400	II	4	Bc2	1,015	1,423	Bc2
<i>Burkholderia xenovorans</i> LB400	III	0				
<i>Ralstonia eutropha</i> H16	II	2	Rs2	395	361	Bc2
<i>Ralstonia eutropha</i> JMP134	II	2	Rs2	398	329	Bc2
<i>Ralstonia metallidurans</i> CH34	II	3	Rs2	402	331	Bc2
<i>Ralstonia solanacearum</i>	II ^d	3	Rs2	404	345	Bc2
<i>Photobacterium profundum</i> SS9	II	7	Vc2	1,703	995	Vc2
<i>Vibrio cholerae</i>	II	9	Vc2	2,110	1,607	Vc2
<i>Vibrio fischeri</i> ES114	II	6	Vc2	1,827	1,154	Vc2
<i>Vibrio parahaemolyticus</i>	II	10 (1)	Vc2 (Bc3)	1,931	1,245	Vc2
<i>Vibrio vulnificus</i> CMCP6	II	16	Vc2	1,948	1,264	Vc2
<i>Vibrio vulnificus</i> YJ016	II	16 (1)	Vc2 (Rs2)	1,948	1,264	Vc2
<i>Rhodobacter sphaeroides</i>	II	0				
<i>Agrobacterium tumefaciens</i> Cereon	Linear	0		1,078	499	Bs2
<i>Agrobacterium tumefaciens</i> UWash	Linear	0		1,083	499	Bs2
<i>Deinococcus radiodurans</i>	II	0				
<i>Haloarcula marismortui</i>	II	0				
<i>Leptospira borgpetersenii</i> JB197	II	0				
<i>Leptospira borgpetersenii</i> L550	II	0				
<i>Leptospira interrogans</i>	II	0				
<i>Leptospira interrogans</i> Copenhageni						
<i>Leptospira interrogans</i> Lai	II	0				
<i>Pseudoalteromonas haloplanktis</i>	II	0		391		Vc2

^a Values in parentheses indicate. The number of sites predicted by the consensus sequence shown in parentheses in column 4.

^b Partitioning site consensus matrices/partitioning proteins used as queries in these searches: Bs2, *Brucella suis* chromosome II; Vc2, *Vibrio cholerae* chromosome II; Rs2, *Ralstonia solanacearum* chromosome II; Bc2, *Burkholderia cenocepacia* chromosome II; Bc3, *Burkholderia cenocepacia* chromosome III.

^c BLAST scores of ParA/RepA or ParB/RepB homologues.

^d Annotated as pGMI1000M.

parS2 sequences found in different families (Fig. 1B). This observation is consistent with the idea that secondary chromosomes have significantly more diverse evolutionary histories than primary chromosomes. Some similarities, however, were observed among sites encoded by secondary replicons within the same or related families (Fig. 1B). For example, all *Burkholderia* chromosome II and chromosome III and *Ralstonia* chromosome II *parS* sites conform to the sequence 5'-TTN(4)CGN(4)AA-3'. Interestingly, no putative *parS3* sites were predicted on the third chromosome of *B. xenovorans*, in contrast to chromosome III of the four other sequenced *Burkholderia* spp., in which six to seven putative sites/chromosome were identified (Table 1). These findings suggest that the third chromosome of *Burkholderia xenovorans* may have a distinct evolutionary lineage from that of its counterparts in other *Burkholderia* strains.

There was significant variability in the number of putative *parS* sites identified on the secondary chromosomes in different

strains (Fig. 3B). All of the secondary chromosomes in *Vibrio* species contain 6 to 16 putative *parS* sites, whereas the secondary chromosomes in *Brucella* and *Ralstonia* species contain only 2 to 3 predicted *parS* sites (Table 1). As was observed with *parS1* sites, most putative secondary chromosome *parS* sites are located in close proximity to the *oriC*s of their respective replicons (Fig. 4B).

A key observation from the searches using the five *parS2*/*parS3* consensus matrices was that, with the exception of one previously described *parS2* site on *V. cholerae* chromosome I (60), no putative *parS2* or *parS3* sites were found on more than one replicon in a single strain. Thus, even though putative *parS2* or *parS3* sites were identified on primary chromosomes, they were not detected on primary chromosomes in strains that contained these sites on secondary chromosomes. This finding, coupled with our observation that no strain contains putative *parS1* sites on more than one rep-

licon, is consistent with the idea that bacteria require replicon-specific *par* loci.

All secondary chromosomes that contain putative *parS*/IR sites also encode putative Par/Rep protein homologues. To identify ParA and ParB homologues on secondary chromosomes, the ParA and ParB proteins encoded by the second chromosomes of *V. cholerae* (Vc2), *B. cenocepacia* (Bc2), and *B. suis* (Bs2, RepA, and RepB) were each compared by BLAST to all proteins in the NCBI database. Since prior studies suggest that Par proteins encoded on secondary chromosomes are more diverse than those encoded on primary chromosomes, we set the minimum BLAST score to 200 for these analyses. Consistent with the observations of Dubarry et al. (13), our findings suggest that Par/Rep proteins encoded on secondary chromosomes, in contrast to ParA and ParB proteins encoded on primary chromosomes, cluster in distinct phylogenetic groups (Table 1) and are more closely related to plasmid-borne partitioning proteins than to those encoded on primary chromosomes. Homologues of both Vc2 ParA and ParB were identified on the secondary chromosomes of all *Vibrio* strains and on plasmids in 10 strains of *Yersinia* sp., *Salmonella* sp., and *Shigella* sp. but not on any primary chromosomes. Homologues of Bc2 ParA and ParB were identified on chromosome II of all 13 *Burkholderia* and *Ralstonia* strains and on 4 of the 5 third chromosomes of *Burkholderia* sp. strains. Interestingly, the only *Burkholderia* sp. third chromosome not predicted to encode ParA and ParB homologues was that of *Burkholderia xenovorans*, the only third chromosome that was not predicted to encode a *parS3* site (Table 1). As shown in Table 1, the Bc2 ParA and ParB are much better conserved on *Burkholderia* second chromosomes than on the third chromosomes of *Burkholderia* sp. or on the second chromosomes of *Ralstonia* sp. (Table 1). Homologues of both Bs2 RepA and RepB were found on the secondary chromosomes of all 4 *Brucella* sp., on the linear chromosomes of both *Agrobacterium tumefaciens* strains, and on a number of plasmids carried by *Alphaproteobacteria* strains. No primary chromosomes or other secondary chromosomes were found to encode homologues of either Bs2 RepA or RepB.

Overall we found a perfect correlation between the presence of *parAB/repAB* genes and cognate *parS*/IR sites on the secondary chromosomes of *Vibrio*, *Burkholderia*, *Ralstonia*, and *Brucella* genera. No putative *parS* sites or Bc2, Vc2, or Bs2 ParB/RepB homologues were identified on the secondary chromosomes of *Deinococcus radiodurans*, *Haloarcula marismortui*, *Pseudoalteromonas haloplanktis*, or *Leptospira* sp., suggesting that if these replicons encode *par* loci, they are probably not related to those encoded by secondary chromosomes in *Burkholderia*, *Brucella*, *Ralstonia*, or *Vibrio* species.

Conclusions. Although the consensus sequence matrix we used to search for *parS* sites was derived from two gram-positive species, putative *parS* sites were identified on the primary chromosomes of 69% of strains from all branches of prokaryotes, and these sites exhibited a high degree of sequence conservation (Fig. 1 and 2). We found that, for the most part, strains that do not contain *parS* sites cluster among relatively few branches of the prokaryotic evolutionary tree. This suggests that *parS* sites (along with Par proteins) evolved very early in the evolution of prokaryotic chromosomes and that the absence of *parS*, *parA*, and/or *parB* in certain strains

likely reflects a loss of one or more of these loci in several ancestral species much later in prokaryotic evolution. The near identity of *parS* sequences among diverse prokaryotic classes is remarkable; binding sites for many other conserved DNA binding proteins, such as LexA and Fur, do not exhibit such a high level of conservation in structure and/or primary sequence (2, 10, 47, 56, 59).

Currently, knowledge of the function of *par* loci on primary chromosomes is rudimentary. Several studies have implicated a role for *par* loci in origin localization and segregation (15, 33, 34, 51), in the separation of sister origins (33), and in the regulation of replication (33, 46, 57). Our finding that the vast majority of *parS* sites are found in the origin-proximal region of the chromosome strongly suggests that the primary function of *par* loci pertains to this part of the chromosome and that this function is highly conserved among diverse bacteria. The biological function, if any, of the small minority of putative *parS* sites found in origin-distal regions of the chromosome remains to be deciphered. We found significant variability among the number of *parS* sites per chromosome: 47 species contain only one putative *parS* site, while 48 species contain six or more putative sites. The functional significance of this variability awaits future exploration. Since we found that 25% of all strains are Par0 and that *par* genes and *parS1* sites can be deleted from several bacteria (16, 21, 29, 34, 51, 60), it is likely that, in most cases, *par* loci encoded on primary chromosomes will not prove to be absolutely required for any essential process, including chromosome partitioning.

Conservation of *parS* sites among diverse genera was not observed in secondary chromosomes, suggesting that the evolutionary lineages of secondary chromosomes, like those of plasmids, are much more diverse than those of primary chromosomes. Unlike the *V. cholerae* chromosome I-encoded Par system (16), the *par* loci encoded on *V. cholerae* chromosome II and on several plasmids are required for the faithful partitioning of these replicons (61). It is possible that *par* loci encoded on other secondary chromosomes will also prove essential for the partitioning of these replicons.

The absence of *parS1* sites from secondary chromosomes and plasmids and the absence (with one exception) of *parS2* sites from primary chromosomes suggest an important difference between prokaryotes and eukaryotes. Eukaryotes can utilize a single mitotic apparatus to mediate the segregation of multiple replicons. Assuming that *par* loci generally play roles in chromosome segregation, our findings suggest that prokaryotes mediate segregation of multiple replicons by utilizing multiple segregation systems. Experimental support for this idea has come from studies of *V. cholerae* and *B. cenocepacia*, where replicon-specific *par*-mediated segregation systems have been described (13, 60). We speculate that the specificity of *parS1* and *parS2* sequences to their respective replicons reflects a mechanism to avoid partitioning incompatibility, as has been observed in cells harboring plasmids that contain similar *parS* sites (7).

Our findings point to some obvious areas for future experimental work. For example, it will be interesting to explore whether ParBS species such as *Staphylococcus aureus* and *Streptococcus pyogenes* contain a functional orthologue of ParA or whether the ParB/*parS* in these species function independently of ParA. Additionally, it should be relatively straight-

forward to test whether the presence of *parS1* sites on secondary chromosomes is indeed deleterious to their segregation. Finally, exploration of segregation by the secondary chromosomes on which no *parS* sites or Par protein homologues were identified may yield novel information regarding chromosome segregation, especially since our findings support the notion that bacterial species with multiple chromosomes require distinct genes and sites to mediate chromosome segregation.

ACKNOWLEDGMENTS

We thank David Rudner, Sarah McLeod, and Brigid Davis for helpful comments on the manuscript.

We acknowledge the support of NIH and HHMI. J.L. was supported by NIH T32-AI07329.

REFERENCES

- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Baichoo, N., T. Wang, R. Ye, and J. D. Helmann. 2002. Global analysis of the *Bacillus subtilis* Fur regulon and the iron starvation stimulon. *Mol. Microbiol.* **45**:1613–1629.
- Barilla, D., M. F. Rosenberg, U. Nobbmann, and F. Hayes. 2005. Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *EMBO J.* **24**:1453–1464.
- Bartosik, A. A., K. Lasocki, J. Mierzejewska, C. M. Thomas, and G. Jagura-Burdzy. 2004. ParB of *Pseudomonas aeruginosa*: interactions with its partner ParA and its target *parS* and specific effects on bacterial growth. *J. Bacteriol.* **186**:6983–6998.
- Bartosik, D., M. Szymanik, and E. Wysocka. 2001. Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J. Bacteriol.* **183**:6234–6243.
- Biek, D. P., and J. Shi. 1994. A single 43-bp *sopC* repeat of plasmid mini-F is sufficient to allow assembly of a functional nucleoprotein partition complex. *Proc. Natl. Acad. Sci. USA* **91**:8027–8031.
- Bouet, J. Y., K. Nordstrom, and D. Lane. 2007. Plasmid partition and incompatibility—the focus shifts. *Mol. Microbiol.* **65**:1405–1414.
- Breier, A. M., and A. D. Grossman. 2007. Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol. Microbiol.* **64**:703–718.
- Calcutt, M. J., and F. J. Schmidt. 1992. Conserved gene arrangement in the origin region of the *Streptomyces coelicolor* chromosome. *J. Bacteriol.* **174**:3220–3226.
- Campoy, S., G. Mazon, A. R. Fernandez de Henestrosa, M. Llagostera, P. B. Monteiro, and J. Barbe. 2002. A new regulatory DNA motif of the gamma subclass Proteobacteria: identification of the LexA protein binding site of the plant pathogen *Xylella fastidiosa*. *Microbiology* **148**:3583–3597.
- Cervin, M. A., G. B. Spiegelman, B. Raether, K. Ohlsen, M. Perego, and J. A. Hoch. 1998. A negative regulator linking chromosome segregation to developmental transcription in *Bacillus subtilis*. *Mol. Microbiol.* **29**:85–95.
- Chai, Y., and S. C. Winans. 2005. RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between *repA* and *repB* for plasmid partitioning and autorepression. *Mol. Microbiol.* **58**:1114–1129.
- Dubarry, N., F. Pasta, and D. Lane. 2006. ParABS systems of the four replicons of *Burkholderia cenocepacia*: new chromosome centromeres confer partition specificity. *J. Bacteriol.* **188**:1489–1496.
- Ebersbach, G., and K. Gerdes. 2005. Plasmid segregation mechanisms. *Annu. Rev. Genet.* **39**:453–479.
- Egan, E. S., M. A. Fogel, and M. K. Waldor. 2005. Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol. Microbiol.* **56**:1129–1138.
- Fogel, M. A., and M. K. Waldor. 2006. A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev.* **20**:3269–3282.
- Friedman, S. A., and S. J. Austin. 1988. The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon. *Plasmid* **19**:103–112.
- Fujita, M. Q., H. Yoshikawa, and N. Ogasawara. 1989. Structure of the *dnaA* region of *Pseudomonas putida*: conservation among three bacteria, *Bacillus subtilis*, *Escherichia coli* and *P. putida*. *Mol. Gen. Genet.* **215**:381–387.
- Garner, E. C., C. S. Campbell, and R. D. Mullins. 2004. Dynamic instability in a DNA-segregating prokaryotic actin homolog. *Science* **306**:1021–1025.
- Garner, E. C., C. S. Campbell, D. B. Weibel, and R. D. Mullins. 2007. Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. *Science* **315**:1270–1274.
- Gerdes, K., J. Moller-Jensen, and R. Bugge Jensen. 2000. Plasmid and chromosome partitioning: surprises from phylogeny. *Mol. Microbiol.* **37**:455–466.
- Godfrin-Estevenson, A. M., F. Pasta, and D. Lane. 2002. The *parAB* gene products of *Pseudomonas putida* exhibit partition activity in both *P. putida* and *Escherichia coli*. *Mol. Microbiol.* **43**:39–49.
- Hassan, A. K., S. Moriya, P. Baumann, H. Yoshikawa, and N. Ogasawara. 1996. Structure of the *dnaA* region of the endosymbiont, *Buchnera aphidicola*, of aphid *Schizaphis graminum*. *DNA Res.* **3**:415–419.
- Hatano, T., Y. Yamaichi, and H. Niki. 2007. Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Mol. Microbiol.* **64**:1198–1213.
- Hayes, F., and S. Austin. 1994. Topological scanning of the P1 plasmid partition site. *J. Mol. Biol.* **243**:190–198.
- Hayes, F., and D. Barilla. 2006. The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat. Rev. Microbiol.* **4**:133–143.
- Hilbert, H., R. Himmelreich, H. Plagens, and R. Herrmann. 1996. Sequence analysis of 56 kb from the genome of the bacterium *Mycoplasm pneumoniae* comprising the *dnaA* region, the *atp* operon and a cluster of ribosomal protein genes. *Nucleic Acids Res.* **24**:628–639.
- Ireton, K., N. W. T. Gunther, and A. D. Grossman. 1994. *spo0J* is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis*. *J. Bacteriol.* **176**:5320–5329.
- Kahng, L. S., and L. Shapiro. 2003. Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J. Bacteriol.* **185**:3384–3391.
- Kim, H. J., M. J. Calcutt, F. J. Schmidt, and K. F. Chater. 2000. Partitioning of the linear chromosome during sporulation of *Streptomyces coelicolor* A3(2) involves an *oriC*-linked *parAB* locus. *J. Bacteriol.* **182**:1313–1320.
- Kolsto, A. B. 1999. Time for a fresh look at the bacterial chromosome. *Trends Microbiol.* **7**:223–226.
- Lasocki, K., A. A. Bartosik, J. Mierzejewska, C. M. Thomas, and G. Jagura-Burdzy. 2007. Deletion of the *parA* (*soj*) homologue in *Pseudomonas aeruginosa* causes ParB instability and affects growth rate, chromosome segregation, and motility. *J. Bacteriol.* **189**:5762–5772.
- Lee, M. J., C. H. Liu, S. Y. Wang, C. T. Huang, and H. Huang. 2006. Characterization of the *Soj/Spo0J* chromosome segregation proteins and identification of putative *parS* sequences in *Helicobacter pylori*. *Biochem. Biophys. Res. Commun.* **342**:744–750.
- Lee, P. S., and A. D. Grossman. 2006. The chromosome partitioning proteins *Soj* (ParA) and *Spo0J* (ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in *Bacillus subtilis*. *Mol. Microbiol.* **60**:853–869.
- Lee, P. S., D. C. Lin, S. Moriya, and A. D. Grossman. 2003. Effects of the chromosome partitioning protein *Spo0J* (ParB) on *oriC* positioning and replication initiation in *Bacillus subtilis*. *J. Bacteriol.* **185**:1326–1337.
- Leonard, T. A., J. Moller-Jensen, and J. Lowe. 2005. Towards understanding the molecular basis of bacterial DNA segregation. *Philos. Trans. R. Soc. Lond. B* **360**:523–535.
- Lim, G. E., A. I. Derman, and J. Pogliano. 2005. Bacterial DNA segregation by dynamic SopA polymers. *Proc. Natl. Acad. Sci. USA* **102**:17658–17663.
- Lin, D. C., and A. D. Grossman. 1998. Identification and characterization of a bacterial chromosome partitioning site. *Cell* **92**:675–685.
- Livny, J., and M. K. Waldor. 2007. Identification of small RNAs in diverse bacterial species. *Curr. Opin. Microbiol.* **10**:96–101.
- Lobry, J. R. 1996. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* **13**:660–665.
- Macke, T. J., D. J. Ecker, R. R. Gutell, D. Gautheret, D. A. Case, and R. Sampath. 2001. RNAMotif, an RNA secondary structure definition and search algorithm. *Nucleic Acids Res.* **29**:4724–4735.
- Mohl, D. A., J. J. Easter, and J. W. Gober. 2001. The chromosome partitioning protein, ParB, is required for cytokinesis in *Caulobacter crescentus*. *Mol. Microbiol.* **42**:741–755.
- Mohl, D. A., and J. W. Gober. 1997. Cell cycle-dependent polar localization of chromosome partitioning proteins in *Caulobacter crescentus*. *Cell* **88**:675–684.
- Moller-Jensen, J., R. B. Jensen, J. Lowe, and K. Gerdes. 2002. Prokaryotic DNA segregation by an actin-like filament. *EMBO J.* **21**:3119–3127.
- Ogasawara, N., and H. Yoshikawa. 1992. Genes and their organization in the replication origin region of the bacterial chromosome. *Mol. Microbiol.* **6**:629–634.
- Ogura, Y., N. Ogasawara, E. J. Harry, and S. Moriya. 2003. Increasing the ratio of *Soj* to *Spo0J* promotes replication initiation in *Bacillus subtilis*. *J. Bacteriol.* **185**:6316–6324.
- Panina, E. M., A. A. Mironov, and M. S. Gelfand. 2001. Comparative analysis of FUR regulons in gamma-proteobacteria. *Nucleic Acids Res.* **29**:5195–5206.
- Qin, M. H., M. V. Madiraju, S. Zachariah, and M. Rajagopalan. 1997. Characterization of the *oriC* region of *Mycobacterium smegmatis*. *J. Bacteriol.* **179**:6311–6317.
- Quisel, J. D., and A. D. Grossman. 2000. Control of sporulation gene expression in *Bacillus subtilis* by the chromosome partitioning proteins *Soj* (ParA) and *Spo0J* (ParB). *J. Bacteriol.* **182**:3446–3451.
- Roulet, E., I. Fisch, T. Junier, P. Bucher, and N. Mermod. 1998. Evaluation

- of computer tools for the prediction of transcription factor binding sites on genomic DNA. In *Silico Biol.* **1**:21–28.
51. Saint-Dic, D., B. P. Frushour, J. H. Kehrl, and L. S. Kahng. 2006. A *parA* homolog selectively influences positioning of the large chromosome origin in *Vibrio cholerae*. *J. Bacteriol.* **188**:5626–5631.
 52. Sharpe, M. E., and J. Errington. 1996. The *Bacillus subtilis* *soj-spo0J* locus is required for a centromere-like function involved in prespore chromosome partitioning. *Mol. Microbiol.* **21**:501–509.
 53. Thanbichler, M., and L. Shapiro. 2006. MipZ, a spatial regulator coordinating chromosome segregation with cell division in *Caulobacter*. *Cell* **126**:147–162.
 54. van Helden, J. 2003. Regulatory sequence analysis tools. *Nucleic Acids Res.* **31**:3593–3596.
 55. von Mering, C., L. J. Jensen, M. Kuhn, S. Chaffron, T. Doerks, B. Kruger, B. Snel, and P. Bork. 2007. STRING 7: recent developments in the integration and prediction of protein interactions. *Nucleic Acids Res.* **35**:D358–D362.
 56. Walker, G. C. 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Rev.* **48**:60–93.
 57. Webb, C. D., A. Teleman, S. Gordon, A. Straight, A. Belmont, D. C. Lin, A. D. Grossman, A. Wright, and R. Losick. 1997. Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* **88**:667–674.
 58. Williams, D. R., D. P. Macartney, and C. M. Thomas. 1998. The partitioning activity of the RK2 central control region requires only *incC*, *korB* and *KorB*-binding site *O_{B3}* but other *KorB*-binding sites form destabilizing complexes in the absence of *O_{B3}*. *Microbiology* **144**:3369–3378.
 59. Winterling, K. W., D. Chafin, J. J. Hayes, J. Sun, A. S. Levine, R. E. Yasbin, and R. Woodgate. 1998. The *Bacillus subtilis* *DinR* binding site: redefinition of the consensus sequence. *J. Bacteriol.* **180**:2201–2211.
 60. Yamaichi, Y., M. A. Fogel, S. M. McLeod, M. P. Hui, and M. K. Waldor. 2007. Distinct centromere-like *parS* sites on the two chromosomes of *Vibrio* species. *J. Bacteriol.* **189**:5314–5324.
 61. Yamaichi, Y., M. A. Fogel, and M. K. Waldor. 2007. *par* genes and the pathology of chromosome loss in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. USA* **104**:630–635.
 62. Yamaichi, Y., and H. Niki. 2000. Active segregation by the *Bacillus subtilis* partitioning system in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **97**:14656–14661.
 63. Yen, M. R., N. T. Lin, C. H. Hung, K. T. Choy, S. F. Weng, and Y. H. Tseng. 2002. *oriC* region and replication termination site, *dif*, of the *Xanthomonas campestris* pv. *campestris* 17 chromosome. *Appl. Environ. Microbiol.* **68**:2924–2933.